(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 301 021 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
09.04.2003 Bulletin 2003/15

(51) Int Cl.7: H04N 1/00, H04N 5/232

(21) Application number: 02078937.6

(22) Date of filing: 23.09.2002

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 04.10.2001 US 971346

(71) Applicant: EASTMAN KODAK COMPANY
Rochester, New York 14650 (US)

(72) Inventors:
- Miller, Michael Eugene,
c/o Eastman Kodak Company
Rochester, New York 14650-2201 (US)
- Jones, Paul W., c/o Eastman Kodak Company
Rochester, New York 14650-2201 (US)
- Yang, Jian, c/o Eastman Kodak Company
Rochester, New York 14650-2201 (US)
- Rabbani, Majid, c/o Eastman Kodak Company
Rochester, New York 14650-2201 (US)

(74) Representative: Parent, Yves et al
KODAK INDUSTRIE,
Département Brevets,
CRT - Zone Industrielle
71102 Chalon-sur-Saône Cedex (FR)

(54) Method and system for displaying an image

(57) A method and system for displaying an image, includes steps and means for: storing image data in a manner that enables retrieval of different spatial regions of an image at different fidelities; determining an viewer's point of gaze on a display; retrieving image data for each spatial region of an image at a fidelity that is a decreasing function of the distance of the regions from the point of gaze; and displaying the retrieved image data on the display;

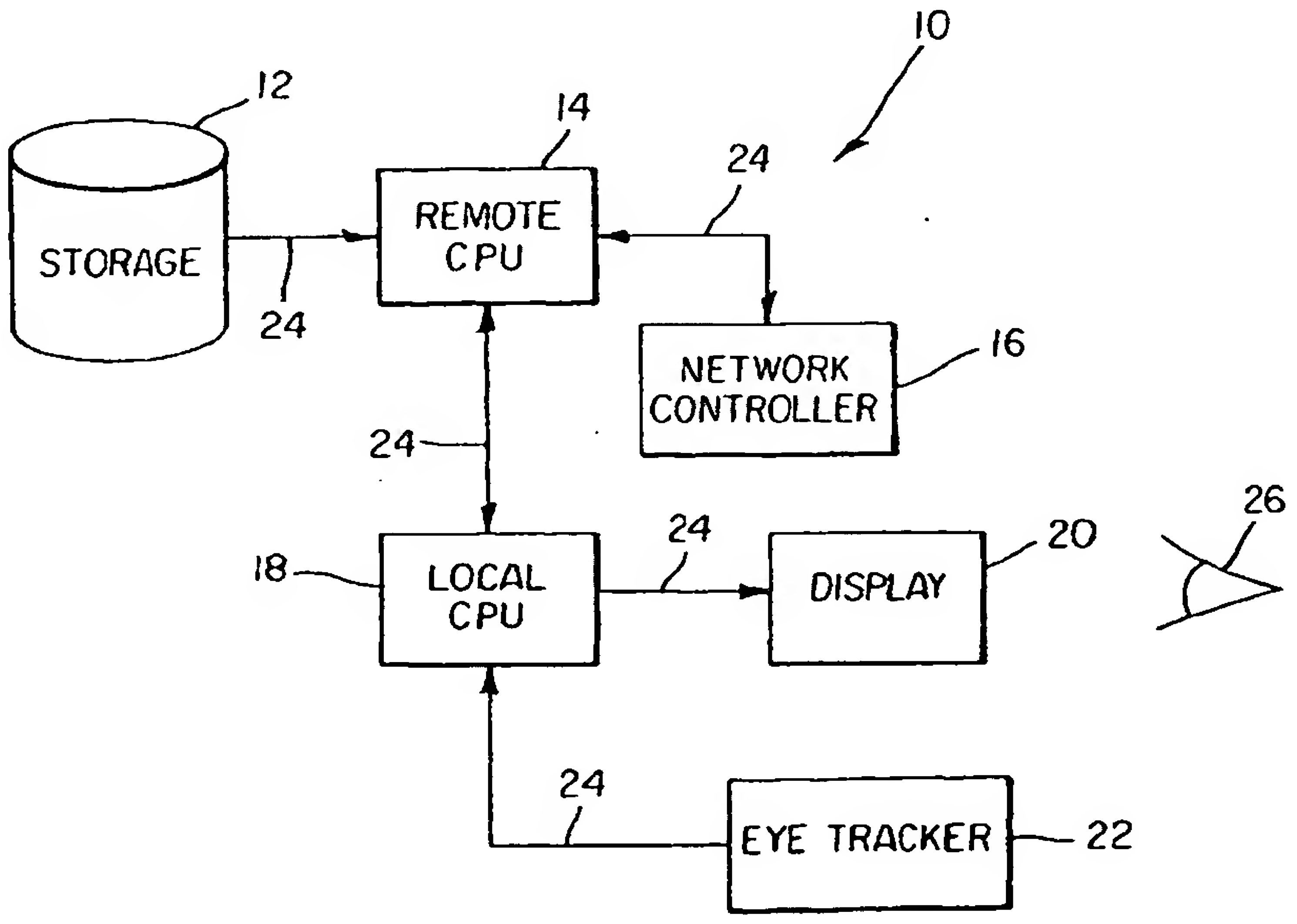

FIG. 1

Printed by Jouve, 75001 PARIS (FR)

storing image data in a manner that enables retrieval of different spatial regions of an image at different fidelities; determining an viewer's point of gaze on a display; retrieving image data for each spatial region of an image at a fidelity that is a decreasing function of the distance of the regions from the point of gaze: and
displaying the retrieved image data on the display.

[0011] The present invention has the advantage that it allows a full resolution image to be stored in such a manner that allows efficient retrieval and transmission of image data that varies in fidelity as a function of the distance from an viewer's point of gaze, reducing system bandwidth requirements for retrieval and transmission. Additionally, the data format allows images to be retrieved and transmitted in a way that allows the necessary image data to be displayed to more than one viewer. Finally, the invention allows the system to react to changes in the accuracy of the eye tracking device and/or changes in system bandwidth in order to deliver an acceptable image to the viewer.

Fig. 1 is a schematic diagram of an image display system according to the present invention;
Fig. 2 is a flow chart summarizing the process used in the present invention;
Fig. 3 is diagram illustrating the relationship between the data structures used in the present invention;
Fig. 4 is a diagram illustrating the code stream used in the present invention;
Fig. 5 is a diagram illustrating the minimum distance of any point in the precinct from a gaze point; and
Fig. 6 is a diagram illustrating one precinct ordering where the ordering of the precincts at a single resolution level progresses from 0 to 15, where the precinct labeled 0 is given the highest priority and the precinct labeled 15 is given the lowest priority.

[0012] Fig. 1 illustrates a system configuration useful in practicing the present invention. The system **10** includes an image storage device **12**, a remote CPU **14**, a network controller **16**, a local CPU **18**, a display **20**, and one or more devices **22** capable of determining the viewer's point of gaze on the display **20**. The system components are connected by a communications network **24**. In this system, the image storage device **12** could be any digital or optical storage device that could store the image information to be viewed. The remote CPU **14** is used primarily to determine the information required from storage based on gaze information and system bandwidth, to retrieve the relevant, compressed information from the storage device, and to transmit this information across the network **24**. The network **24** could be any transmission channel, including digital cable, Internet, or wireless connection. The local CPU **18** receives the visual information, decompresses this information, processes the information for display onto the display **20**, receives information from the eye tracking device, processes it, and transmits gaze information to the remote CPU **14**. The display **20** could be any visual display, but is preferably an immersive display having a field of view of at least X degrees vertical and Y degrees horizontal. The eye tracker **22** could be any device that can be used to monitor the gaze point of a viewer **26**, but ideally it is a system that monitors both head position and gaze position of the viewer.

[0013] It should be recognized that in an alternative embodiment, the remote CPU **14**, network controller **16**, and network **24** could be removed from the system and their functions be performed by local CPU **18**. That is, local CPU **18** is connected directly to the image storage device **12** and retrieves the relevant visual information from the storage device and modifies the characteristics of the imagery that is obtained in response to the retrieval time of the storage device. In this configuration, the primary advantage of the system is to allow apparently very high resolution and field of view imagery to be retrieved from a storage media with relatively low read access time.

[0014] In a preferred embodiment, the system **10** dynamically determines the bandwidth of the transmission channel that is available to the viewer, a number of important system characteristics, and viewer preferences to determine which information is to be retrieved from storage and transmitted to the display device. Fig. 2, provides a flow diagram illustrating a method for determining the image information that needs to be retrieved from the storage device and transmitted to the display device. As shown in this figure, when the system is initiated it first determines **28** the bandwidth that is available for image transmission. This value establishes a limit for the maximum amount of information that can be retrieved from storage and transmitted to the display.

[0015] The system then determines **30** important display and environmental characteristics. During this step, the system may determine important system parameters such as the resolution of the display device, the size of the display device, the viewer's viewing distance, the tone scale and maximum luminance of the display device and important ambient environment variables, such as the level of ambient illumination. As will be shown later, each of these parameters have a direct influence on how the system will retrieve and display information.

[0016] The system also determines **32** the accuracy of the eye tracking device. Although many parameters might be determined, the most important is the expected accuracy of the gaze point. This could be a static value assigned to the eye tracker or may be dynamic, depending on feedback from the system regarding the accuracy of the gaze point calibration that is performed by the eye tracking device.

[0017] Next the system determines **34** if the viewer wishes to view video or still imagery. This distinction is important, primarily because of the tradeoff function between system parameters such as resolution, field of view, and frame rate

[0027] During typical viewing. an viewer makes two different types of eye movements. The most typical of these are discrete eye movements, which are characterized by a period of between 16 and 600 ms during which the point of gaze does not change appreciably. After this phase of the eye movement is completed, a rapid shift in the point of gaze is made before the point of gaze is fixed for another discrete time period of about 16 to 600 ms. This discrete class of eye movements are typified by a relatively constant eye movement velocity of about 500 degrees per second, a value that is almost independent of the amount of change in gaze position (Land. et al., "The Relations Between Head and Eye Movements During Driving," Vision In Vehicles - V., Gale et al. Editors, 1996 Elsevier Science B.V.). Since the eye movement velocity is practically constant regardless of the size of the eye movement, more time will pass between fixations for larger eye movements which typically land in image areas that were previously displayed with very low fidelity. The second type of eye movement is a smooth pursuit eye movement in which the eye continuously follows a moving element in the scene These eye movements are continuous in nature and typically have a velocity of only a few degrees per second. This type of eye movement will result in small changes in the point of gaze where the image was previously displayed with relatively high fidelity.

[0028] In one embodiment, the local processor **18** may simply report the viewers' current point of gaze to the remote processor **14.** In another, more-preferred embodiment. the gaze point estimation provided to the remote processor 14 is based upon an estimate of the eye gaze position at the time the remote processor **14** will deliver data to the local processor **18.** This embodiment is particularly desirable in video systems as it can be estimated that the data for the next image will be delivered after a known time delay.

[0029] To provide this estimate, the current and recent eye movement velocity and acceleration is analyzed to estimate the point of gaze at the time the next frame of data will be delivered. Within this embodiment, it is assumed that excursion of a discrete eye movement will follow a straight line. At any given time t, the location of the point of gaze is determined from an estimate of the velocity and acceleration of the movement in the point of gaze. This determination will preferably be made at a high temporal frequency that is significantly higher than is required for the transmission of image data. That is while image data may be refreshed between 30 and 100 times a second, the eye position will be determined at a frequency that is an order of magnitude higher than this frequency. The velocity and acceleration of the eye movement is determined by calculating the average first and second derivatives of the eye position from a series of the previous gaze points. The change in position of gaze for a time t plus a time delay t_d is then determined using a typical geometric formula such as:

$$D_d = vt_d+at_d^2 \quad (1)$$

Where D_d is the projected distance of travel for the position of gaze, v is the velocity of the eye movement, and a is the acceleration (or deceleration) of the eye movement. This distance, together with the direction of the eye movement, is used to determine the estimate of the gaze position at a time t plus t_d. This estimated position is transmitted to the server to indicate the projected position of the point of gaze when the data will be available at the remote processor. This same approach may be used for smooth pursuit eye movements. However, for smooth pursuit eye movements, the projected distance of the change in the point of gaze will be much smaller.

[0030] Using this method, the region of the image with the highest fidelity will be close to the final point of gaze once an viewer makes a constant point of gaze. This prediction is particularly important when the user makes relatively large eye movements that may require 40 ms or more to execute and which result in a point of gaze in an area that has a very low fidelity before the viewer begins his or her eye movement. Further, this estimate may be continually updated to the remote processor **14.** Since the method described herein, allows transmission of larger spatial extent, low-fidelity image data, followed by smaller spatial extent, high-fidelity image data, refinements of a point of gaze will allow the highest fidelity information to be selected and transmitted very close to the time of display, providing minimal errors in point of gaze estimates.

[0031] To retrieve the appropriate image information as a function of gaze position and the distance from the current gaze position, it is necessary to define a method for determining the image fidelity requirements. This can be accomplished by using a model for certain response characteristics of the human visual system. Specifically, in a preferred embodiment we describe human visual performance using the contrast threshold function, which is a function that specifies the minimum contrast necessary to detect a spatial, sine-wave grating with a spatial frequency *f*.

[0032] It is important to recognize that the spatial resolution of the human eye is inhomogeneous as a function of the distance from the point of gaze. The maximum resolution is found in the fovea, which corresponds to the point of gaze. and resolution decreases as the distance from the fovea is increased. The distance from the center of the point of gaze is also referred to as eccentricity. In an eye-tracked display system, an object is to deliver the highest image fidelity to the viewer's point of gaze and to reduce the fidelity gradually as a function of the eccentricity. Therefore, a model for the contrast threshold function includes a dependence on the distance from the point of gaze. We denote

$$f = \frac{f_p}{\tan^{-1}\left(\frac{s}{n_p d}\right)} \qquad (6)$$

where *s* is the active size of the display along some dimension, n_p is the total number of displayed pixels along the same dimension, and *d* is the viewing distance. As the viewing distance is increased, a given value of f_p will map to a larger value of *f*, which leads to a decrease in the contrast threshold value produced by Eqs. 2 and 3.

[0037] Values such as the active area of the display, number of display pixels, and display reflectance may be recorded in the display's memory during manufacture and reported to the system using industry standard protocols, such as VESA's data display channel. Minimum and maximum luminance may be determined based on sensors that are designed to measure the luminance of the display or they may be derived from other relevant display parameters, such as the beam current in a CRT. Ambient illumination can be measured through the use of a light sensor attached to a display and the viewer's viewing distance may be derived from the apparatus that is used to determine head and eye gaze position. While these display and environmental variables may be provided by the described means, it is possible that some or all of this information may be unavailable. In such cases, it is necessary to assume a nominal value for each of the parameters that were discussed.

[0038] The image storage device **12** contains high-resolution information for all spatial locations in an image, as it is not known a priori where the point of gaze will be for a given individual and a given image or image sequence. For a practical and cost-effective system, an efficient compressed representation is required for the high-resolution images in order to minimize the amount of information that must be stored. Any number of well-known compression techniques, such as the current JPEG or MPEG standards, can be used to provide this efficient storage. However, the present invention places additional constraints on the compressed representation in that it must allow for the rapid retrieval of spatial and resolution subsets of the high-resolution image information as the gaze point changes. Moreover, these spatial and resolution subsets must be compactly represented so that the necessary image information can be transmitted across the network within allotted bandwidth. The efficient retrieval and transmission of spatial and resolution subsets is very difficult to accomplish with the current JPEG and MPEG standards. They are primarily designed to provide constant resolution across the full spatial extent of an image, which is inconsistent with the concept of foveated imaging.

[0039] There are other compressed data representations that are better suited for use in a foveated imaging system. In particular, the JPEG 2000 compression standard has recently been defined, and JPEG 2000 provides a framework that integrates very well with the requirements of foveated imaging. This is because JPEG 2000 uses a wavelet transform as a key component in the compression process. A wavelet transform decomposes an N x N original image into an N x N set of wavelet coefficients, where each coefficient corresponds both to a given spatial location in the original image and to a given range of frequencies (called a subband). Thus, the wavelet coefficients provide a space-frequency representation, which allows convenient access to the spatial and resolution subsets that are needed in foveated imaging.

[0040] JPEG 2000 is not the only compression technique that uses a wavelet decomposition (or more generally, a subband decomposition), and any other technique that uses a subband decomposition would provide similar benefits. Furthermore, resolution-based hierarchical decompositions (e.g., a Gaussian pyramid) can be used to provide access to spatial and resolution subsets, although these representations are generally less efficient than a wavelet or subband representation. However, in the following description, the JPEG 2000 standard is used because it is well-defined and contains appropriate data structures to enable foveated imaging. It is understood that similar concepts can be used with other wavelet, subband, or resolution-based hierarchical compression techniques, and in fact, it may be advantageous to deviate from the JPEG 2000 standard to provide features that are not enabled with a fully compliant JPEG 2000 system. JPEG 2000 is primarily a standard for still-frame images, but it can easily be applied to each frame in an image sequence. In the present invention, it is assumed that an image sequence is represented as a set of independently encoded frames. While this may result in lower compression efficiency than a technique that takes advantage of frame-to-frame correlation (such as MPEG), it greatly simplifies access to the necessary data as the gaze position or system bandwidth requirements change over time.

[0041] To understand the use of JPEG 2000 in the present invention, it is first necessary to review some of the compressed data structures that are provided within the standard. These data structures include: components, tiles, resolution levels, precincts, and layers. All of these data structures relate to the organization of the wavelet coefficients within the compressed codestream. The various data structures provide: (1) access to color channels, e.g., RGB or YCbCr, (through components); (2) access to spatial regions (through tiles); (3) access to frequency regions (through resolutions levels); (4) access to space-frequency regions (through precincts); and (5) access to coefficient amplitudes clature. this type of progression ordering is known as a "resolution level-layer-component-position" progression. This ordering is also used when storing the low-resolution data on the storage device so that it is a simple matter to stream the corresponding contiguous data packets onto the network.

[0049] Given this background image that represents the low-resolution information, it is then necessary to fill in higher resolution detail information in accordance with viewer's point of gaze and the corresponding contrast threshold function values across the field of view. This requirement suggests that the remaining compressed data packets should be ordered first according to spatial location and then according to resolution levels and finally according to precincts and layers. In addition, the data can be organized according to the color component, so that component information can be easily prioritized (for example, sending the luminance channel first because of its perceptual importance). In JPEG 2000 terminology, this type of ordering is known as a "component-position-resolution level-layer" ordering. The codestream is thus organized so that all data packets for a given tile are contiguous, and within a tile, all data packets for a given precinct are contiguous. In this way, the data packets for a particular spatial location can be efficiently accessed by locating the corresponding tile and/or precinct boundaries in the codestream. JPEG 2000 provides unique data "markers" that allow one to easily determine the tile positions in the codestream, but it may be advantageous to specify a separate table of byte-count offsets. This table consists of spatial locations in the original image (e.g., the point of gaze) and for each spatial location, there is a corresponding byte-count offset (e.g., from the beginning of the file) that indicates the start of the packets for a given tile and/or precinct. This type of lookup table provides an efficient means to locate the desired data packets when forming the codestream for transmission over the network.

[0050] The basic ordering of the codestream is depicted in Fig. 4. The first section **74** is organized according to the "resolution level-layer-component-position" ordering to allow efficient streaming of the low-resolution background information. The second section **76** is organized according to the "component-position-resolution level-layer" to allow efficient streaming of the higher resolution data for a particular image region. Unique marker segments **78, 80** in JPEG 2000 are used to indicate which ordering is being used at a given point in the codestream so that the codestream may be correctly interpreted.

[0051] It is important to make a distinction between the codestream that is stored on the storage device and the codestream that is transmitted over the network. As mentioned previously, the stored codestream contains high-resolution information for all spatial locations in an image, while the transmitted codestream is a subset of the stored codestream in accordance with the gaze point and bandwidth constraints. Although the general structure of both codestreams follows that shown in Fig. 4, the specific ordering of the data packets for the high-resolution detail information will be different because of a need to prioritize the data so that the fidelity in the gaze point region of the viewer is improved first. If this prioritization is not performed. it is possible that there may not be sufficient bandwidth to provide the desired level of fidelity in the gaze point region, i.e., too much of the available bandwidth may have been consumed in representing less critical areas away from the point of gaze. This means that the data packets representing high-resolution detail information for the gaze point region must be retrieved from storage and placed at the beginning of the transmitted codestream. Subsequent data packets in the codestream would correspond to the detail information for areas away from the point of gaze.

[0052] The prioritization of the data packets is performed using the distance *r* from the center of the gaze position, which is provided by the eye tracker **22** to both the local and remote CPUs **18** and **14** respectively. For each precinct in an image, the minimum distance of any point in the precinct from the gaze point is computed, as shown in Fig. 5. Data packets that represent the precincts (over one or more resolution levels) are prioritized the order of closest distance to furthest distance. In this way, the fidelity is increased first in the gaze point region **82,** and the surrounding regions are then refined subsequently. An example of the precinct ordering is illustrated in Fig. 6, where the ordering of the precincts **68** at a single resolution level progresses from 0 to 15, where the precinct labeled 0 is given the highest priority and the precinct labeled 15 is given the lowest priority. This same prioritization could also be performed using the larger spatial structure of tiles, instead of precincts, which would provide less localization, but may be more efficient in terms of streaming compressed data from the server. It is possible to send the compressed data for the precincts and/or tiles using a minimum of overhead information because the gaze point **82** (and consequently the ordering) is known at both the local and remote CPUs **18** and **14** respectively.

[0053] Now, the distance *r* from the gaze point only tells us how the data packets from the various precincts should be prioritized; it does not indicate how much information should be sent for each precinct. Because of the bandwidth constraints of the network, the goal is to send only as much detail as is needed for each spatial region in the image. The determination of the amount of detail information for each spatial region is performed using the contrast threshold function that was described previously. The contrast threshold function can be used to compute the precision that is required for the wavelet coefficients to ensure that an viewer will not be able to detect any degradations in the displayed image.

[0054] To apply the contrast threshold function to the wavelet coefficients, it is necessary to understand the impact of the bit plane encoding process that is used in JPEG 2000. Suppose a wavelet coefficient is initially quantized with a step size of Δ. If the *k* least significant bit planes of the coefficient are then discarded, the effective quantizer step gaze points. As before we have the issues of: 1) prioritization of the data packets and 2) precision of the wavelet coefficients.

[0061] Regarding the data packet prioritization, it can generally be assumed that the various viewers have equal importance. Thus, a reasonable strategy is to alternate data packets in the codestream among the viewers. For example, a precinct for the gaze point of viewer 1 would be transmitted first, followed a precinct for the gaze point of viewer 2. Then, the next precinct for viewer 1 would be sent, followed by the next precinct for viewer 2, and so on. If it is known that one viewer has a greater importance, the data packets for more precincts could be sent first for that viewer. An example of this scenario is when there are two viewers, and the eye tracker is able to detect that one viewer is frequently closing his or her eyes (i.e., falling asleep).

[0062] Now, we address the precision of the wavelet coefficients. Although the methods described by Eqs. 7 and 8 are still valid, the fields of view for the different viewers will overlap, and we must consider the most critical viewing condition at each spatial location in the image. For a given precinct, we must compute its distance *r* from the gaze point of each viewer, and then use the closest gaze point distance in computing the quantizer step size via Eq. 7. As a result, even though the prioritization of a precinct may depend upon one viewer, then precision that is used for the precinct may be determined by another viewer.

[0063] It must also be recognized that the bandwidth constraints may impose limitations that prevent the viewer from obtaining satisfactory fidelity. At this time, the viewer may desire to make other tradeoffs, including reducing the field of view of the image and/or the frame rate of video imagery. Here the viewer simply indicates the desired field of view of the image and/or the frame rate of the image through a dialog. These selections will affect the bandwidth that is available for a given image, thus potentially allowing for improved picture fidelity, depending upon the specific selections.

[0064] When field of view (image size) is changed, the image data outside the selected field of view can be truncated, and no information is transmitted for the corresponding tiles or precincts. If the frame rate is altered, the system simply transmits and displays the images at the selected rate. If a slower frame rate is selected, the fidelity of each frame will increase, at the possible expense of non-smooth motion. A higher frame rate will provide smoother motion, but at the expense of lower fidelity for each frame. The impact of these tradeoffs depends greatly upon the scene content.

[0065] An eye tracking device plays an important role in the previously described embodiment of the system. However, it may not always be practical to dynamically determine the gaze positions of all viewers of a system. When active gaze point estimation is not practical, an alternative means for determining the likely gaze positions within an image may be used to replace the eye tracking device. In this embodiment of the system, two different techniques may be applied to determine the likely gaze positions in an image. These include performing image analysis to determine likely positions of gaze, and measurement and statistical description of probability density maps for the typical points of gaze within an image as determined by a set of viewers.

[0066] It should be noted, however, that neither of these techniques are deterministic as any viewer's point of gaze changes about every 100 to 300 ms. For this reason, neither technique can be used to predict an viewer's exact point of gaze. Instead, each of these techniques can only predict the regions of the image where an viewer is most likely to direct his or her point of gaze during normal image viewing. To achieve robust application of either of these techniques, the image processing system must therefore enable the image to be processed to accommodate multiple gaze positions.

[0067] The application of image analysis to predict areas of an image that are likely to attract an viewer's gaze typically involves determining the image regions that contain one or more particularly high contrast edges or other salient information. Methods for determining probability maps through image analysis are well documented in the art and descriptions can be found by several authors, including: Itti et al., "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research,* vol. 40, 2000, pp. 1489-1506; Cartier et al., "Target attractiveness model for field-of-view search," *Optical Engineering,* 1998, vol 37(7), pp. 1923-1936; Reinagel et al., "Natural scene statistics at the centre of gaze," *Computational Neural Systems,* 1999, vol. 10, pp. 341-350.

[0068] To determine a probability density map of likely points of gaze for a group of typical viewers, it is necessary to have them perform a task similar to the task of the final viewers. Each viewer views the image on a representative display while active eye tracking is being performed and the viewer's gaze positions are recorded. The data from this group of viewers is then combined into a single data structure containing coordinates for each of the points of gaze determined for each individual frame of image data.

[0069] It should also be noted that as described earlier, the local processor transmits the estimated gaze position to the remote processor. As this information is transmitted, the system can store this data as a function of the image that is displayed. Therefore, if a baseline system is built that includes an eye tracking system, this system can be used to archive gaze information from a potentially large number of viewers. This data may be used to determine probability density maps for systems that do not include eye tracking devices. This same data may also be leveraged for many other uses, including the advertising or price determination of real estate within the imagery (e.g., a person who purchases an advertisement slot or a bill board within the virtual environment might be charged based upon the number gaze from the image content.

12. The method claimed in claim 1, wherein the point of gaze of an viewer is determined by measuring the points of gaze of a plurality of viewers of an image and determining the point of gaze as a function of the measured points.

13. The method claimed in claim 1, further comprising the steps of compressing the image data prior to storage and decompressing the retrieved image data prior to display.

14. The method claimed in claim 13, wherein the compressed image data is compressed using wavelet based compression.

15. The method claimed in claim 13, wherein the compressed image data is compressed using Gaussian pyramid based compression.

16. A method for processing a digital image, comprising the steps of:

a) determining a plurality of likely points of gaze on the image; and
b) processing the digital image such that each spatial region of the image is represented at a fidelity that is a decreasing function of the distance of each region from the likely points of gaze.

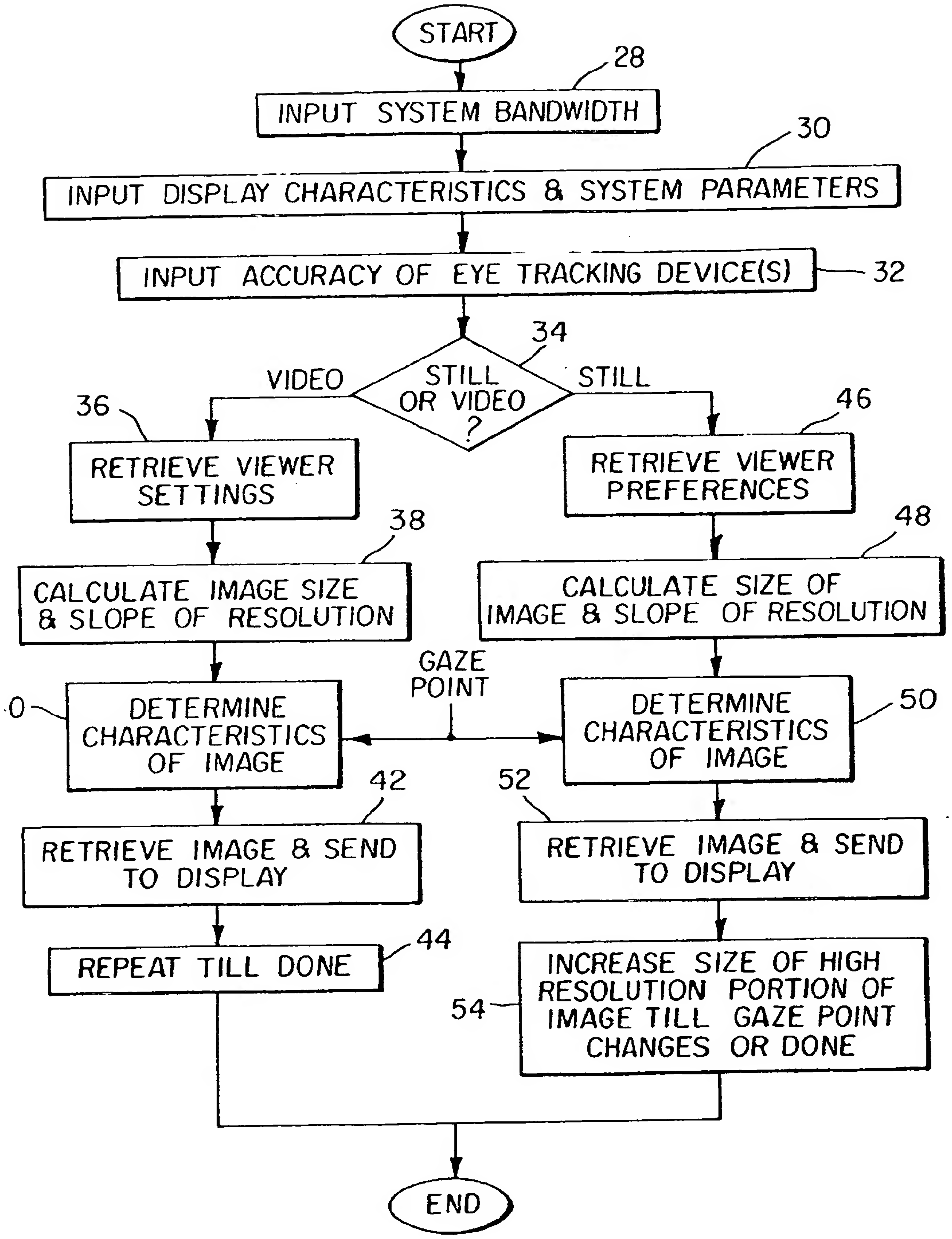
START
28
INPUT SYSTEM BANDWIDTH
30
INPUT DISPLAY CHARACTERISTICS & SYSTEM PARAMETERS
INPUT ACCURACY OF EYE TRACKING DEVICE(S)
32
34
VIDEO
STILL OR VIDEO ?
STILL
36
RETRIEVE VIEWER SETTINGS
46
RETRIEVE VIEWER PREFERENCES
38
CALCULATE IMAGE SIZE & SLOPE OF RESOLUTION
48
CALCULATE SIZE OF IMAGE & SLOPE OF RESOLUTION
GAZE POINT
0
DETERMINE CHARACTERISTICS OF IMAGE
50
DETERMINE CHARACTERISTICS OF IMAGE
42
RETRIEVE IMAGE & SEND TO DISPLAY
52
RETRIEVE IMAGE & SEND TO DISPLAY
44
REPEAT TILL DONE
54
INCREASE SIZE OF HIGH RESOLUTION PORTION OF IMAGE TILL GAZE POINT CHANGES OR DONE
END

FIG. 2

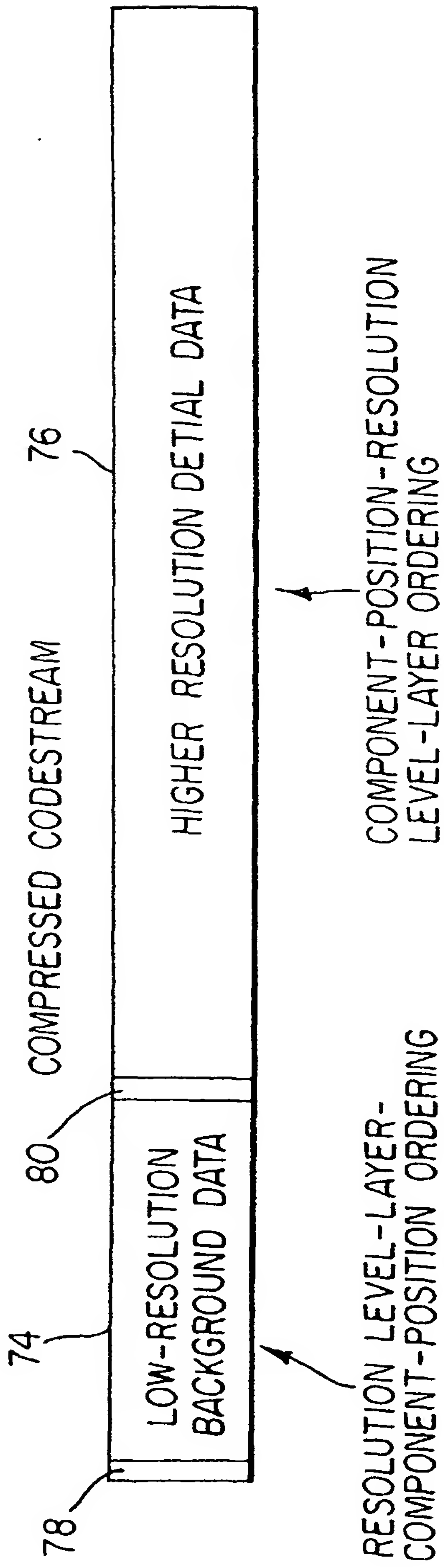
COMPRESSED CODESTREAM
76
HIGHER RESOLUTION DETIAL DATA
80
74
LOW-RESOLUTION BACKGROUND DATA
78
COMPONENT-POSITION-RESOLUTION LEVEL-LAYER ORDERING
RESOLUTION LEVEL-LAYER-COMPONENT-POSITION ORDERING

FIG. 4